



# Baseline groundwater monitoring for shale gas extraction: definition of baseline conditions and recommendations from a real site (Wysin, Northern Poland)

N. Montcoudiol<sup>1</sup> · D. Banks<sup>1</sup> · C. Isherwood<sup>2</sup> · A. Gunning<sup>2</sup> · N. Burnside<sup>1</sup>

Received: 14 June 2018 / Accepted: 28 January 2019 / Published online: 2 February 2019  
© The Author(s) 2019

## Abstract

Public concerns have been raised regarding the use of hydraulic fracturing for shale gas extraction and its potential impact on the environment. The absence of baseline monitoring data in the US experience has been identified as a major issue. Here, results from a 6-month groundwater baseline monitoring study near an active shale gas pad in northern Poland are presented. The data collected in dedicated boreholes include groundwater samples analysed for inorganic constituents, dissolved gases and stable isotopes ( $\delta^2\text{H}$  and  $\delta^{18}\text{O}$ ) and downhole temperature and conductivity measurements at 15-min intervals. A robust statistical analysis combined with an estimation of data uncertainty helps to identify spatial and temporal variability within the datasets. As a result, baseline conditions are defined using confidence intervals around the mean on a per-well basis and these will serve for future reference for this site. The groundwater chemical composition is similar to regional background levels and typical of Quaternary aquifers in the region. It is also consistent with previous baseline monitoring carried out by the Polish Geological Institute. Only manganese and bromide occur in groundwater at concentrations above Polish drinking water standards. Based on this work, the paper provides some recommendations for future baseline monitoring and identifies areas for future research such as use of statistics for high-frequency datasets.

**Keywords** Hydraulic fracturing · Shale gas · Europe · Baseline monitoring · Groundwater · Statistical analysis

## Introduction

Exploitation of shale gas by hydraulic fracturing ('fracking') gained its controversial status after allegations that many drinking water wells in the vicinity of US shale gas sites had suffered detrimental changes in abstracted groundwater quality (Howarth et al. 2011; Schnoor 2012). Systematic reviews have largely failed to substantiate the most extreme claims, although there are a few situations in the published literature where well integrity failure has been identified or

is strongly suspected to have been the cause of impairment to groundwater quality (Bair et al. 2010; Darrah et al. 2014; Jackson et al. 2013a; US EPA 2015). The lack of baseline data in the USA led to difficulties in identifying the existence, magnitude and cause of alleged groundwater quality changes (Vidic et al. 2013). Nevertheless, uncritical reporting of groundwater contamination allegations triggered public opposition in Europe at the time the first shale gas drilling permits were awarded (Williams et al. 2015). In 2011, hydraulic fracturing at Preese Hall 1 near Blackpool, the first UK shale gas well, was suspended after a 2.3-magnitude induced earthquake (Huw et al. 2014). This event drew further public attention to the shale gas industry and its use of hydraulic fracturing. As a consequence, a more cautious approach has been taken in Europe as compared to the USA. Several potential impacts have been identified (Mair et al. 2012), and the need for baseline monitoring prior to any shale gas activity has been widely acknowledged by the scientific community (Jackson et al. 2013b; Mair et al. 2012). Baseline monitoring will allow the detection of changes and trigger corrective actions from the operator if necessary or,

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s11600-019-00254-w>) contains supplementary material, which is available to authorized users.

---

✉ N. Montcoudiol  
[Nelly.Montcoudiol@glasgow.ac.uk](mailto:Nelly.Montcoudiol@glasgow.ac.uk)

<sup>1</sup> School of Engineering, University of Glasgow,  
Glasgow G12 8QQ, Scotland, UK

<sup>2</sup> RSKW, Stirling University Innovation Park,  
Stirling FK9 4NF, Scotland, UK

in the absence of changes, to reassure the public regarding the safety of shale gas operations. To date, very few published baseline studies have been carried out prior to hydraulic fracturing and they were mostly focusing on dissolved methane concentrations (Bell et al. 2017; Humez et al. 2016; Moritz et al. 2015; Schloemer et al. 2016; Siegel et al. 2015)—which is often insufficient for a contamination diagnostic (Lefebvre 2017)—and/or considered baseline conditions at regional scale (Harkness et al. 2017; Rhodes and Horton 2015; Sloto 2014).

In Europe, Poland is one of the countries with the largest estimated reserves (PGI-NRI 2012; US EIA/ARI 2013), located within the Lower Palaeozoic Baltic–Podlasie–Lublin Basins. Poland was the leader of shale gas exploration until the end of 2016 when efforts ceased, largely due to disappointing production rates from test wells (Cantoni 2018). During the period 2010–2016, a total of twenty-eight horizontal wells were hydraulically stimulated (Pyssa 2017), of which seven sites were independently monitored by the Polish Geological Institute (English acronym PGI-NRI; Koniecznyńska et al. 2015), producing limited baseline data, e.g. one single-event sampling for water bodies, 1.5 months of background seismicity at the Łebień site (Koniecznyńska et al. 2011). One particular site (Wysin in Pomerania, northern Poland, within the Baltic Basin) was subject to an intense international monitoring study within the framework of the EU Horizon 2020 SHEER project. The overall goal of the project was to develop best practice in order to understand, prevent and mitigate environmental impacts and risks associated with shale gas activities. To achieve this objective, the first multidisciplinary continuous monitoring effort at an instrumented shale gas site in Europe was undertaken, where seismicity, air and groundwater monitoring was undertaken before, during and after fracking. Together with subsequent British Geological Survey's (BGS) monitoring efforts at two planned shale gas sites in England (Smedley et al. 2015), these programmes provide the only known examples of long-term baseline monitoring prior to planned shale gas fracking.

Baseline monitoring started progressively from July 2015 (air quality) and by the end of 2015, the entire monitoring network was functional. The baseline monitoring lasted until hydraulic fracturing was carried out in June and July 2016. The duration of operational and post-fracking monitoring varied from a few months (seismicity) to 18 months (groundwater) (more details available in López-Comino et al. 2018). Here, this paper focuses on the results from the baseline groundwater monitoring at site level, using a network of four dedicated boreholes. The objectives of the paper are threefold: (1) the quality of the shallow drinking water aquifer located near the Wysin shale gas pad is assessed using dedicated boreholes. The monitoring results are then compared with the regional background levels and

previous local baseline monitoring to evaluate the suitability of the boreholes for monitoring (Jackson and Heagle 2016); (2) a systematic statistical approach to small and continuous datasets is combined with the estimation of data uncertainties to assess spatial and temporal variability and provide a consistent framework for robust data analysis; (3) unbiased baseline conditions at local scale are defined and could serve as future reference levels to identify potential groundwater contamination from shale gas activities. The paper concludes with some recommendations arising from this work.

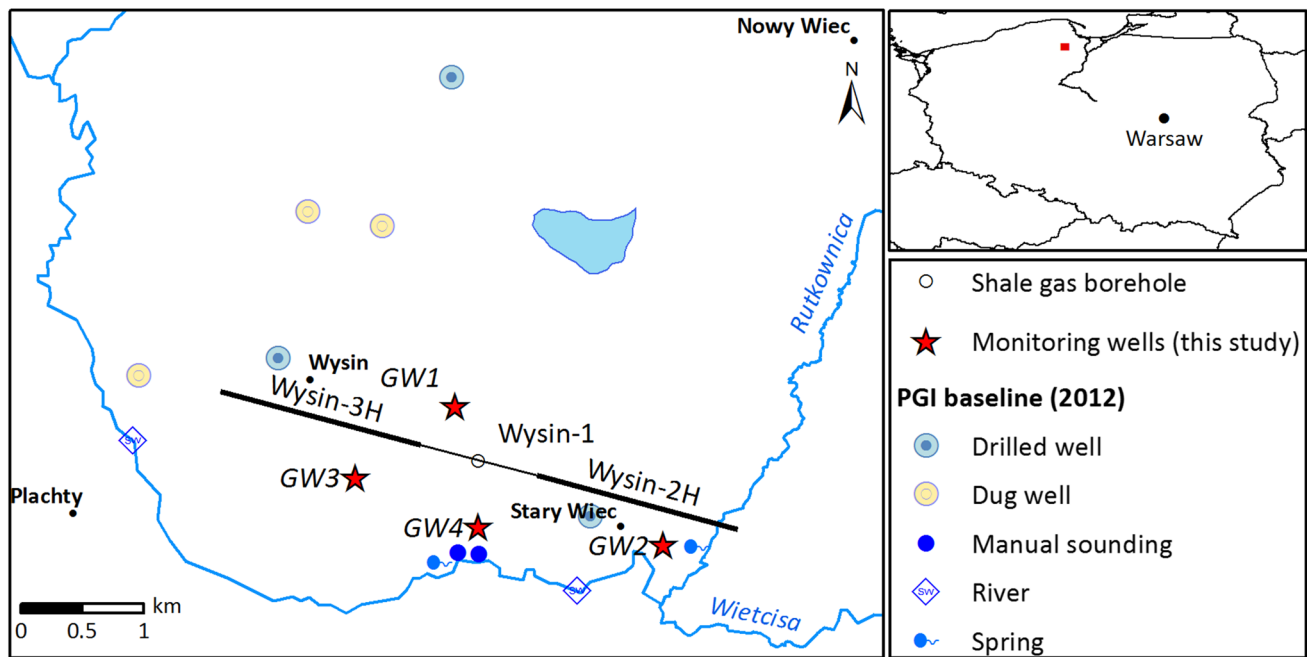
## Study area

The Stara Kiszewa shale gas concession area is located in the Pomerania province of northern Poland, about 50 km south-west of Gdansk city (Fig. 1). It covers an area of 981.5 km<sup>2</sup>. Within the Stara Kiszewa concession area, exploration works targeting hydrocarbons have been carried out since the 1960s. Seismic surveys were carried out in the region during two campaigns in 2004–2008 and 2011–2012 (OGI n.d.).

## Geology of the region

The Pomerania province is located within the Baltic Basin, considered to be one of the most promising shale gas basins in Europe (PGI-NRI 2012; US EIA/ARI 2013). It has a simple geological structure which has experienced relatively little tectonic deformation. The rock sequence runs from the Palaeozoic to the Mesozoic periods. The prospective target rocks for shale gas are the Lower Palaeozoic Middle Silurian Wenlock (typically 400 m thick) and Lower Silurian Llandovery Formations (typically less than 100 m thick), which both contain dark grey to black organic shales that commonly exhibit strong gas shows in exploration wells (Brownfield et al. 2015).

The geology of the sub-surface was confirmed by the drilling of the vertical Wysin-1 borehole for shale gas exploration. The Wysin-1 borehole was drilled in 2013 to a final depth of 4040 m, after penetrating a 54.5-m-thick section of Middle Cambrian rocks (Makos 2014). These rocks comprise black mudstones and clays with interbeds of fine-grained quartz sandstones. They are overlain by 30 m of Ordovician rocks comprising marl, claystone and shale belonging to the Prabuty formation. The Silurian succession is more substantial, about 1900 m in thickness, including 35 m of the Llandovery Formation shales. The rest of the sequence includes in order of succession, about 400 m of Permian, 600 m of Triassic, 300 m of Jurassic and 550 m of Cretaceous strata. The sequence is completed by 235 m of Tertiary and Quaternary sediments (Koniecznyńska et al. 2014, 2015).



**Fig. 1** Location of the shale gas pad, the groundwater monitoring boreholes from this study and the sampling points from the Polish Geological Institute's baseline monitoring (Koniecznyńska et al. 2014)

Subsequent to the initial vertical Wysin-1 exploration well, two additional boreholes, with laterals intended for hydraulic fracturing, were drilled from the same pad (Wysin-2H and Wysin-3H) during autumn 2015. Wysin-2H reaches a true vertical depth (TVD) of 3865 m with a target formation of the Lower Silurian Wenlock Formation strata. Wysin-3H reaches a TVD of 3974 m targeting the Ordovician age strata.

### Hydrogeology of the region

The regional groundwater system is known as the Baltic Artesian Basin (BAB). It forms a complex multi-layered hydrogeological system that covers about 480,000 km<sup>2</sup>. The thicknesses and permeabilities of the successive aquifer and aquitard layers are variable. The main aquifers are present in Cambrian, Upper Ordovician–Upper Silurian, part of Devonian (where present), part of Carboniferous (where present), part of Upper Permian, Lower-Middle Jurassic, Cretaceous, Paleogene, Neogene and Quaternary age strata (Virbulis et al. 2013). The Upper Cretaceous system forms the predominant regional fresh water aquifer (Sadurski 1986). Fresh water is also found in the most recent sediment layers, especially the Pleistocene aquifer, one of the richest groundwater sources in Poland (Jaworska-Szulc 2009; OGI n.d.). The whole system is mostly confined apart from the unconfined/leaky confined Quaternary system. The permeability of the aquitards allows some exchange of groundwater between the

layers of active flow (Jaworska-Szulc 2009; Sadurski 1986; Virbulis et al. 2013).

The shale gas concession is located in a predominantly rural region where the population relies on groundwater for drinking water and irrigation water for agriculture. The hilly landscape's geomorphology is a result of the last glaciations (the presence of the Scandinavian continental glacier), combined with the erosive and accumulative activity of the rivers. Locally, the Quaternary aquifer is considered to be multi-layered and is commonly conceptualised with three recognised aquifers separated by two aquitards (Koniecznyńska et al. 2015; OGI n.d.). The uppermost aquifer is associated with river valleys sediments. This shallow aquifer has only local significance as a potable water supply due both to its limited thickness and deteriorated water quality (Koniecznyńska et al. 2014, 2015). A discontinuous layer of silty clay separates this aquifer from the middle one and offers partial protection from the contaminated superficial aquifer (Koniecznyńska et al. 2015). The middle aquifer forms the main water resource due to its easy accessibility (15–50 m below ground level), its sandy/gravelly composition and its good water quality (OGI n.d.). The third aquifer is largely unexploited due to its depth (given the presence of the second aquifer) and is therefore poorly understood due to scarcity of data (Koniecznyńska et al. 2015), but is thought to be supplied by infiltration from the upper aquifers (Koniecznyńska et al. 2014). In terms of the EU Water Framework Directive, parts of the Gołębiewo Intra-Moraine groundwater body (groundwater body No. 30; EEA 2014)

are located within the shale gas concession, including the area of the drilling pad (Koniecznyńska et al. 2015). However, the groundwater protection area is located outside the concession (OGI n.d.).

About 800 m south of the Wysin shale gas pad lies the Wietcisa River, which flows from west to east, towards the Wisła (Vistula), and acts as a local discharge zone for the Quaternary aquifer. The Rutkownica River, a tributary to the Wietcisa, is located 2 km east of the pad and also acts as a local discharge zone (Fig. 1; Koniecznyńska et al. 2015). Consequently, the main shallow groundwater flow direction is north–south, with lateral flow in the east and west directions. In addition to the rivers, the area is characterised by the presence of lakes, waterlogged areas in endorheic depressions and seasonal or permanent wetlands (OGI n.d.).

## Materials and methods

### Monitoring plan

The monitoring strategy is based on the compilation of existing data, mainly from the Polish Geological Institute, the environmental impact assessment undertaken prior to drilling, and the literature review presented in the introduction.

Prior to the SHEER groundwater monitoring network, four monitoring boreholes were installed in the vicinity of the pad and are believed to be part of the operator's monitoring plan. Unfortunately, access to these data was not granted. The SHEER groundwater monitoring network also consists of four boreholes, whose locations were selected based on a hydrogeological model developed by the PGI-NRI (Koniecznyńska et al. 2014). The final locations were constrained by owner's permissions to access their land. These boreholes were drilled in the main Quaternary aquifer in November 2015 (Fig. 1). Their characteristics are summarised in Table 1 of the Electronic Supplementary Material.

Drilling of the monitoring boreholes suggests that the commonly accepted three-layer conceptual model is significantly oversimplified. The Quaternary multi-layer aquifer is

highly heterogeneous, as is often the case with fluvioglacial sediments, and is characterised by the presence of lower-permeability layers or lenses interbedded within highly permeable layers. Simplified geological cross sections based on these boreholes and other available borehole logs in the vicinity of shale gas pad can be found in Gunning et al. (2017).

The groundwater baseline monitoring started about 1 month after the installation of the monitoring boreholes, in December 2015. The baseline monitoring period lasted 6 months during which four sampling rounds were completed. Emphasis was placed on the characterisation of the aquifer system and its background concentrations in terms of inorganic compounds, dissolved gases and isotopic signatures. Inorganic constituents are comparatively easy to sample and analyse. Their fate within the sub-surface is well known with a few of them behaving conservatively (e.g. chloride). They may be good indicators of surface spills reaching the aquifer, one of the most common sources of contamination at shale gas sites (US EPA 2016). Much less is known about organic compounds (Gordalla et al. 2013; Kahrilas et al. 2015; Luek and Gonsior 2017; Stringfellow et al. 2014).

### Field work

#### Continuous monitoring of temperature and electrical conductivity

Downhole probes (CTD-Divers, Schlumberger) were installed in each borehole in December 2015, at the midpoint of the screened interval (Fig. 1 in ESM). They record absolute pressure (non-vented sensor), temperature and specific electrical conductivity (at 25 °C) at 15-min intervals. In this paper, the focus is on the groundwater quality. Water level data processing, including correction for barometric effects and interpretation, is fully described in Gunning et al. (2017).

Specifications for temperature and conductivity sensors are the same for the four downhole probes. The temperature

**Table 1** List of statistical tests performed on the groundwater quality data (laboratory analyses and downhole measurements)

Test name	Short description	$\alpha$ -level
Shapiro–Wilk test (Shapiro and Wilk 1965)	Test for normality	0.1 ( $n < 10$ ) 0.05 ( $10 \leq n \leq 20$ ) 0.01 ( $n > 20$ )
Rank van Neumann ratio test (Bartels 1982; US EPA 2009)	Statistical independence	0.01
Levene's test (Levene 1961)	Equality of variance	0.01
Kruskal–Wallis test (Kruskal and Wallis 1952)	Spatial variability	0.05
Duplicate method (Grøn et al. 2007; JCGM 2008; Ramsey and Ellison 2007; Witczak et al. 2006)	Uncertainty on measurements	0.05 (95% confidence level)
Confidence intervals	Definition of baseline conditions	0.05 (95% confidence level)

was measured with an accuracy of  $\pm 0.1$  °C and a resolution of 0.01 °C. Accuracy and resolution were, respectively,  $\pm 1\%$  and 0.1% of the reading for the electrical conductivity.

### Groundwater sampling

As part of the baseline monitoring, a total of four sampling campaigns were carried out between December 2015 and June 2016, before the fracking started.

The wells were purged using a submersible pump (GRUNDFOS® model SQE-2-85) placed a few metres below the water level to ensure a good-quality purging (Fig. 1 in ESM). Samples were taken after pumping out three well-bore volumes and once the physico-chemical parameters had stabilised. The physico-chemical parameters (temperature, pH, specific electrical conductivity, dissolved oxygen and oxidation–reduction potential) were measured using a multi-parameter probe (model YSI Professional Plus) which was calibrated each day.

A 1-l plastic bottle was filled up to the neck for major ions, minor and trace element analyses. In addition to the laboratory measurement, total alkalinity was also determined in the field using a HACH® kit. Each titration was performed with 0.16 N sulphuric acid on 100 ml of freshly collected sample, to which a bromocresol green–methyl red indicator powder pillow was added. Samples for dissolved gas analyses were collected in 20-ml brown glass vials filled without bubbles and closed by a septum cap. Following this same sampling procedure, water stable isotope ( $\delta^{18}\text{O}$  and  $\delta^2\text{H}$ ) samples were collected in duplicate or triplicate in 15-ml glass vials filled without bubbles. The bottle cap is taped to minimise losses by evaporation. All samples are kept at 4 °C during storage and transport.

### Laboratory analyses

The cation and metal analytical techniques included (1) inductively coupled plasma optical emission spectrophotometry (ICP-OES) for calcium (Ca), magnesium (Mg), sodium (Na), potassium (K), iron (Fe), aluminium (Al), boron (B) and lithium (Li), and (2) inductively coupled plasma mass spectrometry (ICP-MS) for manganese (Mn), strontium (Sr), arsenic (As), barium (Ba), cadmium (Cd), chromium (Cr), copper (Cu), mercury (Hg), nickel (Ni), lead (Pb), antimony (Sb) and selenium (Se). These two analytical techniques were applied to a sample aliquot that was filtered through a 0.45- $\mu\text{m}$  filter and acidified (nitric acid) in the laboratory prior to analysis. The anions chloride ( $\text{Cl}^-$ ), fluoride ( $\text{F}^-$ ) and bromide ( $\text{Br}^-$ ) were analysed by ion chromatography (IC). Sulphate ( $\text{SO}_4^{2-}$ ) concentrations were determined by photometric analysis (turbidity upon reaction with barium chloride) in a Konelab 30i discrete analyser. In addition to field determination, alkalinity was analysed in the laboratory

by titration on filtered samples. All analyses were carried out by the Concept Life Sciences commercial laboratory at East Kilbride, UK (accredited ISO 17025:2005).

Dissolved gases (methane, ethane, propane and ethene), analysed by the Concept Life Sciences laboratory in Manchester, UK, were undertaken by gas chromatography coupled to a flame ionisation detector (GC-FID).

Water stable isotope analyses were undertaken at the SUERC laboratories in East Kilbride, UK. For  $\delta^{18}\text{O}$  analysis, each sample was over-gassed with a 1%  $\text{CO}_2$ -in-He mixture for 5 min and left to equilibrate for a further 24 h. A sample volume of 2 ml was then analysed using standard techniques on a Thermo Scientific Delta V mass spectrometer set at 25 °C. Final  $\delta^{18}\text{O}$  values were produced using the method established by Nelson (2000). For  $\delta^2\text{H}$  analysis, sample and standard waters are injected directly into a chromium furnace at 800 °C (Donnelly et al. 2001), with the evolved  $\text{H}_2$  gas analysed online via a VG Optima mass spectrometer. Final values for both  $\delta^{18}\text{O}$  and  $\delta^2\text{H}$  are reported as per mil (‰) variations from the Vienna Standard Mean Ocean Water (V-SMOW) in standard delta notation. In-run repeat analyses of water standards (international standards V-SMOW and GISP—Greenland Ice Sheet Precipitation, and internal standard Lt Std) gave a reproducibility better than  $\pm 0.3\text{‰}$  for  $\delta^{18}\text{O}$  and  $\pm 3\text{‰}$  for  $\delta^2\text{H}$ .

### Quality control procedure

#### Duplicates and blanks

To ensure good-quality sampling and laboratory analyses, sequential duplicates represent 10% of the total number of analysed samples (1 duplicate every 10 samples), and field and transport blanks represent 5% (1 of each type for every 20 samples). Blank preparation, field handling and interpretation can be found in Gunning et al. (2017). This procedure was applied for the entire monitoring. As a result, only one pair of duplicates was taken during baseline monitoring (GW1; March 2016). Additional duplicates and blanks were taken during the operational monitoring period.

#### Uncertainties of laboratory and downhole measurements

Measurement uncertainties ( $u$ ) arising from analyses ( $s_{\text{analytical}}^2$ ) and sampling ( $s_{\text{sampling}}^2$ ) were estimated using an empirical approach based on a statistical model (Eq. 1; Ramsey and Ellison 2007). The complexity of the empirical method depends on the number of sampling teams involved in the sampling process and the number of sampling protocols followed. Here, only one sampling team is involved using the same protocol (exceptions detailed below). As a result, the ‘duplicate’ method was applied which allows



estimation of a combined analytical and sampling uncertainty. By this method, only the precision component of each uncertainty ( $s_{a,\text{precision}}^2$  and  $s_{s,\text{precision}}^2$  in Eqs. 2 and 3) is evaluated and the bias ( $s_{a,\text{bias}}^2$  and  $s_{s,\text{bias}}^2$  in Eqs. 2 and 3) component is assumed to be negligible. The theory behind the duplicate method together with examples can be found in Grøn et al. (2007), JCGM (2008), Ramsey and Ellison (2007) and Witczak et al. (2006). More details on the calculation of uncertainties applied to this study are available in Montcoudiol et al. (2018).

$$u = s_{\text{meas}} = \sqrt{s_{\text{sampling}}^2 + s_{\text{analytical}}^2} \quad (1)$$

$$s_{\text{analytical}}^2 = s_{a,\text{precision}}^2 + s_{a,\text{bias}}^2 \quad (2)$$

$$s_{\text{sampling}}^2 = s_{s,\text{precision}}^2 + s_{s,\text{bias}}^2 \quad (3)$$

It is assumed that measurement uncertainties are constant during the entire monitoring programme. Therefore, uncertainties are estimated by using all duplicates collected over the course of the 2-year monitoring programme (representing 10% of the collected samples). A total of five duplicate pairs are available (Table 3 in ESM). Grøn et al. (2007) recommend using a minimum of eight sets of duplicates to obtain a reliable estimate. Fewer can result in an overestimation of the uncertainty (Ramsey and Ellison 2007).

The same method is used for isotope data, for which a large number of duplicates/triplicates is available during the baseline monitoring programme (8 duplicates and 8 triplicates). The ‘duplicate’ method was modified as necessary when applied to triplicate samples as explained in Montcoudiol et al. (2018). The method provides the uncertainty of one measurement. Rules of error propagation were applied to estimate the uncertainty for the average duplicate/triplicate result.

Uncertainty for temperature and conductivity measurements is estimated using the manufacturer’s specifications provided in “Continuous monitoring of temperature and electrical conductivity” section. The resolution is related to the precision of the measurement, i.e. the lowest precision that can be obtained, which results in a precision usually exceeding the resolution. The accuracy relates to the systematic error or bias in the measurement.

### **Ion balance error**

The ion balance error (IBE) was calculated for each analysis. Results of electrical balances are compiled in Table 2 of the ESM. They are mostly between 5 and 10%. Additional discussion regarding these results is available in Gunning et al.

(2017). For duplicates, the sample with the IBE closest to zero was retained for further data analysis and interpretation.

## **Statistical analysis**

A series of statistical tests were performed on the data to define the baseline conditions at the site. These are listed in Table 1, in which  $\alpha$ -levels from the US EPA (2009) are included. Full description of the tests including their underlying assumptions is available in Helsel and Hirsch (2002) and US EPA (2009).

### **Basic assumptions**

Normality, statistical independence and equality of variance are the basic assumptions for a number of tests and are tested first. Independence might be achieved when sampling on an occasional basis, but this is certainly not the case for data collected by the downhole probes at short intervals (e.g. 15-min intervals). Although physical independence does not guarantee statistical independence (but makes it more likely), physical independence of the measurements is first estimated for these datasets by calculating a minimum time between measurements to ensure that distinct volumes of groundwater are measured. This method, based on Darcy’s law, is outlined in Chou (2004) and US EPA (2009). The normality and equality of variance tests are complemented by a graphical method, i.e. boxplots from Tukey (1977).

Spatial variability and temporal variability are tested by using the Kruskal–Wallis test (Kruskal and Wallis 1952) and its reverse version, respectively. Despite the small size of the datasets (3 groups of size 5 or less or 4 or more groups of size 4 or less per group; Helsel and Hirsch 2002), the large-sample approximation is used due to the presence of ties (samples with identical concentrations) within the data, with the exception of strontium for which an exact test is computed (Helsel and Hirsch 2002).

### **Confidence intervals**

Baseline conditions are defined by calculating confidence intervals. Ideally, a nonparametric approach would be used. However, due to the limited amount of data, nonparametric intervals cannot be defined with a sufficient level of confidence (i.e. 87.5% with  $n=4$  vs. 95% or more). Parametric confidence intervals are based on assumptions of independent and identically distributed measurements. In other words, there should be no outliers, the measurements are statistically independent, there is no trend (no temporal variability) and no spatial variability, and the data are approximately normally distributed (US EPA 2009). A parametric approach is therefore used, after testing for normality,

independence and the presence of spatial and temporal variability (Table 1).

## Results and interpretation

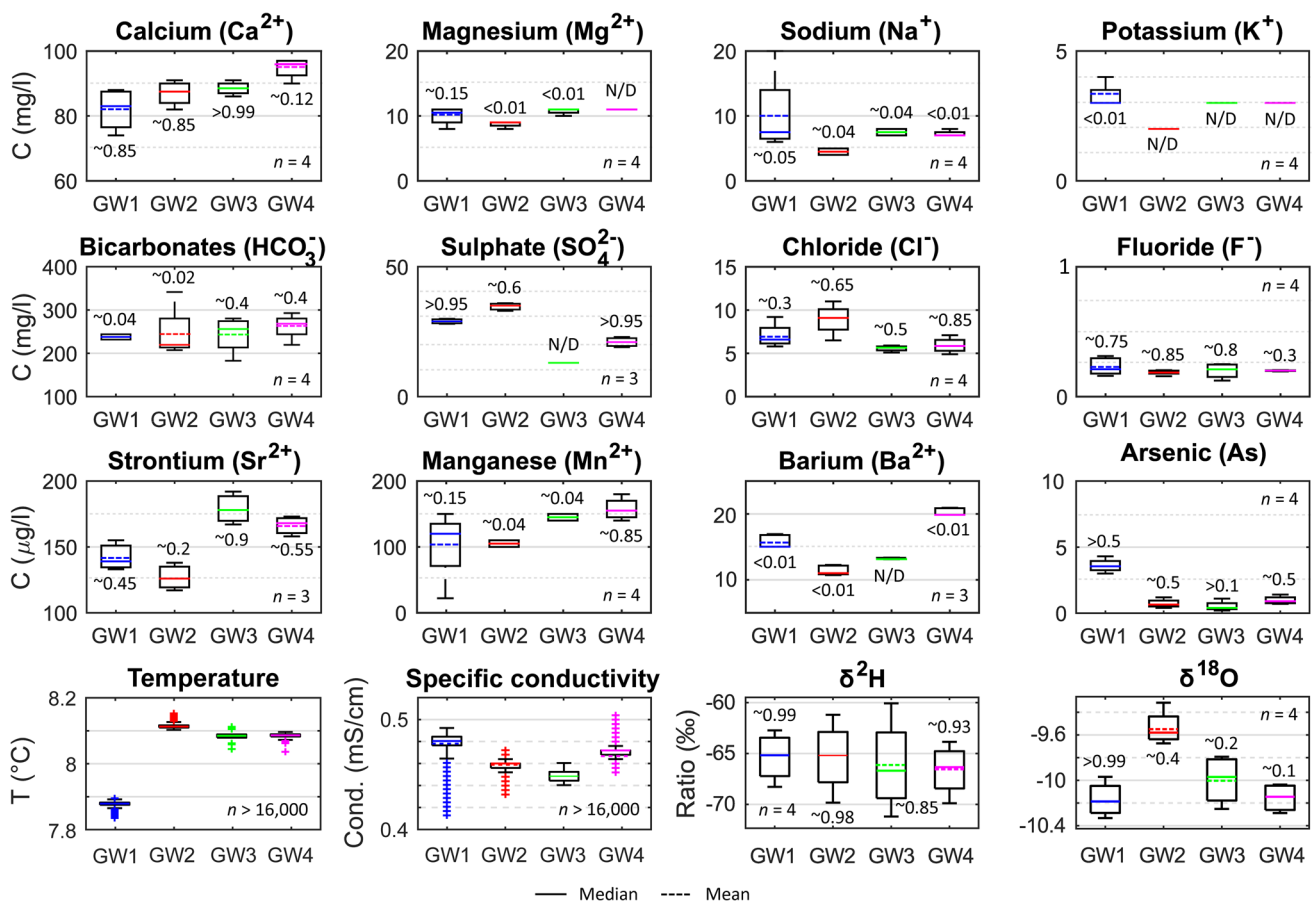
### Results of statistical tests

#### Assessing normality

In general, the mean and the median are similar, but they are not always located in the middle of the boxplot box (representing the interquartile range), suggesting some skew in the data distribution (Fig. 2). Some extreme values are present although not identified as outliers (e.g. Na in GW1,  $\text{HCO}_3^-$  in GW2 and Mn in GW1). The outliers observed for temperature and conductivity data result from the impact of well-purging during groundwater sampling (Figs. 2 and 3 in ESM).

Results from the Shapiro–Wilk test (Shapiro and Wilk 1965) indicate that in most cases the hypothesis of normal distribution cannot be rejected, considering the  $\alpha$ -level of 0.1 recommended by Helsel and Hirsch (2002) and the US EPA (2009) to maximise the utility of the test for small datasets ( $n < 10$ ; Fig. 2). The parameters and wells where this hypothesis is rejected are those exhibiting small concentrations with limited variations (e.g. Mg, Na, K and Ba) or with one sample with a very different concentration (visible in the boxplots in Fig. 2). In the latter case, a departure from normality is plausible. As a result, for inorganic parameters and isotope ratios, the more robust nonparametric tests are used where possible and normality is assumed when a sufficient level of confidence cannot be reached (e.g. confidence intervals).

Normality is not tested for the large datasets obtained with the downhole probes ( $n > 16,000$ ), for which the central limit theorem (Pólya 1920) applies. The departure from normality (tailing) is largely due to the impact of sampling on groundwater conditions and is not representative of the



**Fig. 2** Boxplots for major, minor and trace elements in significant concentrations, temperature, conductivity and water stable isotopes. The bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The maximum whisker length is 1.5 times the

difference between the 25th and 75th percentiles. Points not included within the whiskers are plotted individually as outliers (symbol +). The value near each boxplot is the  $p$  value from the Shapiro–Wilk test. *N/D* not defined (when all concentrations were identical)

background conditions. Independence of data at 15-min intervals is discussed in the following subsection.

### Statistical independence

For inorganic constituents and isotope ratios, statistical independence with respect to temporal variability is confirmed by the results of the rank von Neumann ratio test, for which the  $p$  values are larger than the 0.01  $\alpha$  value. However, no explicit adjustment of the ratio has been developed in the presence of ties among ranks, resulting in very approximate  $p$  values.

Data at 15-min intervals (i.e. temperature and conductivity) are not statistically independent with respect to time. Based on data available prior to drilling of the monitoring boreholes, physical independence is estimated to be obtained within 9 h. Additional data from drilling confirm this first estimate except for GW2 where the local hydraulic conductivity is significantly lower (Table 2). For this borehole, the minimum time would be around 46 h (~2 days).

Results from the rank von Neumann ratio test show that statistical independence is reached over 2-week or 3-week time intervals for most parameters. It must, however, be borne in mind that the  $p$  values in Table 2 are approximate due to the presence of ties and therefore are only indicative. The  $p$  values are also sensitive to the impact of sampling

on the values (i.e. selection of the initial data point from a total of 1920 possibilities), especially for the conductivity in GW1, which shows a longer impact from the sampling (up to 10 days; Fig. 3 in ESM). The new datasets at 3-week intervals are checked again for normality. The hypothesis of normality is rejected for conductivity in GW1 and GW3 with an  $\alpha$ -level of 0.1.

### Equality of variance

The examination of the boxplots suggests significant difference between the variances of each well (Fig. 2) except for arsenic, strontium and isotope ratios. From the Levene's test results, significant departure from equality of variance is observed for potassium and sodium only (Table 3). The validity of the test is questionable for potassium since concentrations are mostly constant over time in each well (Table 2 in ESM). For sodium, the  $p$  value is close to the  $\alpha$ -level. Log-transformation does not improve the results of the test. Therefore, original concentration data are used for the Kruskal–Wallis test. For the other parameters, there is insufficient evidence of significant difference between the variances in the different wells.

For temperature and conductivity data, the boxplots suggest similar variance (Fig. 2). However, because the Kruskal–Wallis test assumes statistical independence

**Table 2** Assessing physical and statistical independence

Physical independence using Darcy's law					
Parameter	Prior information	GW1	GW2	GW3	GW4
Hydraulic conductivity $K$ (m/s)	$1 \times 10^{-4}$	$6.8 \times 10^{-5}$	$1.2 \times 10^{-5}$	$6.5 \times 10^{-5}$	$7.4 \times 10^{-5}$
Hydraulic gradient $i$ (m/m)	$1 \times 10^{-2}$	$8.3 \times 10^{-3}$	$1.3 \times 10^{-2}$	$1.3 \times 10^{-2}$	$2.0 \times 10^{-2}$
Effective porosity $n_e$ (–)	0.2	0.20	0.12	0.14	0.20
Intergranular velocity $v_H$ (m/s)	$5 \times 10^{-6}$	$3 \times 10^{-6}$	$1 \times 10^{-6}$	$6 \times 10^{-6}$	$7 \times 10^{-6}$
Diameter of the well $h$ (m)	0.165				
Minimum time between measurements $t$ (h)	~9	~15	~46	~8	~7
Statistical independence using the rank von Neumann ratio test ( $\alpha$ -level=0.01)					
Test	Parameter	GW1	GW2	GW3	GW4
$t = 2$ d ( $n = 88$ )	Temperature	<0.005	<0.005	<0.005	<0.005
	Conductivity	<0.005	<0.005	<0.005	<0.005
$t = 1$ week ( $n = 26$ )	Temperature	<0.005	<0.005	<0.01	~0.01
	Conductivity	>0.1	<0.005	<0.005	<0.005
$t = 2$ weeks ( $n = 13$ )	Temperature	>0.1	<0.01	>0.01	>0.1
	Conductivity	>0.1	>0.025	<0.005	<0.01
$t = 3$ weeks ( $n = 8$ )	Temperature	>0.1	>0.1	<0.005	>0.1
	Conductivity	>0.1	>0.1	>0.05	<0.005

Prior information comes from shape files transmitted by the Polish Geological Institute. Hydraulic conductivity values result from the interpretation of pumping test data (GT 2015). Hydraulic gradients are estimated based on the results of a preliminary groundwater numerical model (Gunning et al. 2017). Effective porosity values are estimated from granulometric curves (data available in GT 2015). Figures in the bottom part of the table are the  $p$  values from the rank von Neumann ratio test



**Table 3**  $p$  values for the Levene's test (equality of variance) and the Kruskal–Wallis test (spatial and temporal variability)

Parameter	Levene's ( $\alpha=0.01$ )	Spatial Kruskal– Wallis ( $\alpha=0.05$ )	Temporal Kruskal–Wallis ( $\alpha=0.05$ )
Calcium	0.02	0.07	0.12
Magnesium	0.05	0.05	0.88*
Sodium	0.01	0.03	0.93*
Potassium	0.002	0.03	0.96*
Alkalinity	0.12	0.63	0.68
Sulphate	0.20	0.02	0.84*
Chloride	0.33	0.03	0.67*
Strontium	0.88	0.01	0.50*
Manganese	0.02	0.02	0.82*
Barium	0.01	0.02	0.96*
Fluoride	0.11	0.29	0.05
Arsenic	0.88	0.02	0.46*
$\delta^2\text{H}$	0.71	0.86	0.01
$\delta^{18}\text{O}$	0.19	0.02	0.30*
Temperature	0.26	< 0.001	0.90*
Specific conductivity	0.17	0.001	0.50*

\*Temporal variability likely to be overshadowed by the spatial variability

$\alpha$ -levels recommended by the US EPA (2009). Datasets at 3-week intervals are used for temperature and conductivity

between the observations, the entire dataset is not appropriate and the previously defined 3-week interval datasets are used. Results from the Levene's test show  $p$  values above the chosen  $\alpha$ -level (Table 3), and similar variance between the datasets is assumed.

### Spatial and temporal variability

Results from the Kruskal–Wallis test (spatial variability) show that only bicarbonate and fluoride (and possibly calcium) concentrations and  $\delta^2\text{H}$  ratios have  $p$  values higher than the chosen  $\alpha$ -level (Table 3). This means that these parameters are not affected by spatial variability and additional statistical tests could be run on an inter-well basis. For the other parameters, statistical tests on per-well basis are required.

The reverse Kruskal–Wallis test for temporal variability only works for parameters with no significant spatial variability; otherwise, it may overshadow the temporal variability (as is the case in this study; Table 3). Of these four parameters, only the  $\delta^2\text{H}$  isotope ratios show significant temporal variability.

### Uncertainties

Maximum uncertainties on the concentrations (at 95% confidence level) are in the 20–25% range for calcium, chloride, fluoride and manganese (Table 3 in ESM). For comparison, maximum uncertainties in the EU Directive 98/83/EC on the quality of drinking water (EC 1998) are 30% for arsenic and 20% for fluoride (Cl, Mn, Na,  $\text{SO}_4^{2-}$  and conductivity have been removed in the new directive proposal; EC 2018).

Estimates of uncertainty are made more difficult by a change of analytical method for alkalinity determination (from December 2016), affecting the last duplicate. Based on the entire dataset, it is believed that the new technique has a better analytical precision: before December 2016, analytical results appear to be rounded up to the nearest ten, and after December 2016, to the nearest digit. This has some effect on the results of the duplicate method. Table 3 of the Electronic Supplementary Material shows that samples analysed by the initial technique have the same alkalinity values, whereas the last pair of duplicates analysed with the new method presents a difference of 3 mg/l. From the same date, the filtration and acidification for cations and metals was carried out in the field instead of the laboratory. However, this change seems to have had an insignificant impact on the duplicate concentrations.

Uncertainties cannot be defined when all the sets of duplicates have the same concentrations. This is the case for magnesium, potassium and barium. This is due to the precision of the analytical results which have been rounded up to the nearest digit, and suggests that sampling does not add uncertainty larger than the precision of the analysis results. Maximal theoretical uncertainties can be assessed taking into consideration rounded numbers. With the initial analytical technique for alkalinity determination, the concentration uncertainty is  $\pm 5$  mg/l which is not accounted for in the calculation of uncertainty. Taking into account the maximal difference between two duplicates (9 mg/l), the uncertainty could be up to 9.5% (at 95% confidence level). In the same fashion, maximum uncertainty could be 25%, 16%, > 100% and 22% for arsenic ( $\pm 0.05$   $\mu\text{g/l}$ ), barium ( $\pm 0.5$   $\mu\text{g/l}$ ), potassium ( $\pm 0.5$  mg/l) and magnesium concentrations ( $\pm 0.5$  mg/l), respectively.

Uncertainties of  $\delta^2\text{H}$  and  $\delta^{18}\text{O}$  isotope ratios calculated using the 'duplicate' method are very similar to the reproducibility given by the laboratory (Table 4 in ESM). The calculated uncertainties combined analytical and sampling precision, whereas reproducibility only accounts for analytical precision. Results demonstrate that sampling does not add more uncertainty to the isotopic ratios than analysis. Reproducibility data are used as a proxy for uncertainties of  $\delta^2\text{H}$  and  $\delta^{18}\text{O}$  isotope ratios.

For the temperature and conductivity measurements, the precision is assumed to be 10 times the resolution.

Therefore, precision and accuracy are the same:  $\pm 0.1$  °C and  $\pm 1\%$  of the reading for temperature and conductivity, respectively. Combined, the total uncertainty of temperature and conductivity is  $\pm 0.14$  °C and  $\pm 1.4\%$  of the reading, respectively (at 95% confidence level).

### Chemical characterisation of the aquifer

All the data handling and processing have been extensively discussed in Gunning et al. (2017). Only the main results are presented in this section.

#### Major, minor and trace elements

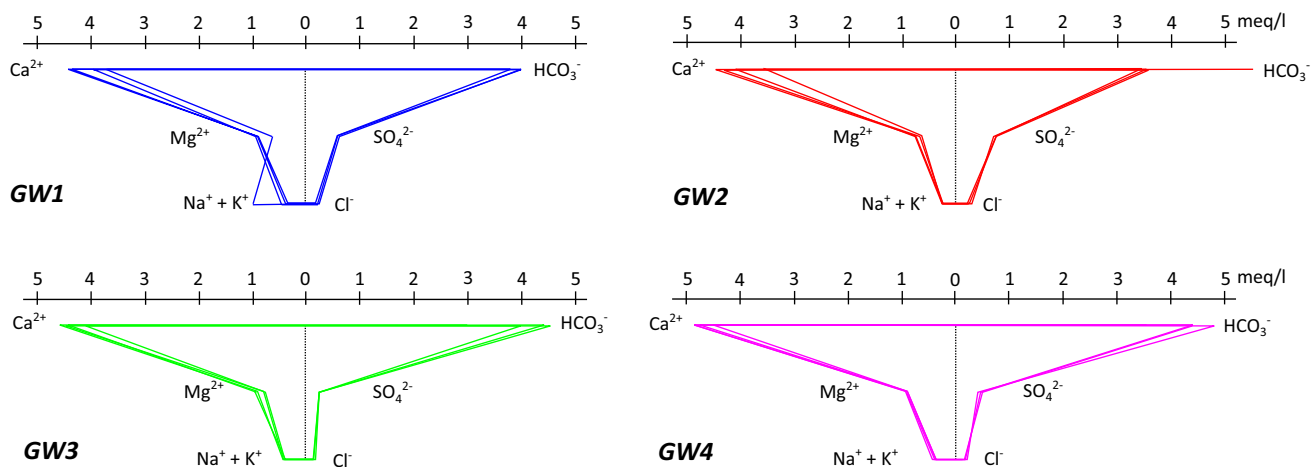
The groundwater is characterised by low mineralisation and is of Ca-HCO<sub>3</sub> water type (Fig. 3). Low mineralisation is also corroborated by the measurements of groundwater conductivity, varying within the range of 440–500  $\mu\text{S}/\text{cm}$  (Fig. 3 in ESM). Generally speaking, the groundwater quality is similar in the four monitoring boreholes, i.e. concentrations and conductivity values are of the same order of magnitude, suggesting good hydraulic connectivity between them. From a statistical point of view (Kruskal–Wallis test,  $\alpha = 0.05$ ), only bicarbonates, fluoride and possibly calcium concentrations are similar in the four wells (Table 3), whereas for the other parameters the concentrations in at least one well differ significantly from the other wells.

When considering the uncertainties in concentrations, the groundwater chemistry shows limited temporal variability during the baseline monitoring (Fig. 4). The only exception would be bicarbonate and sulphate concentrations, which appear to be significantly different from one sampling event to another. This is not statistically

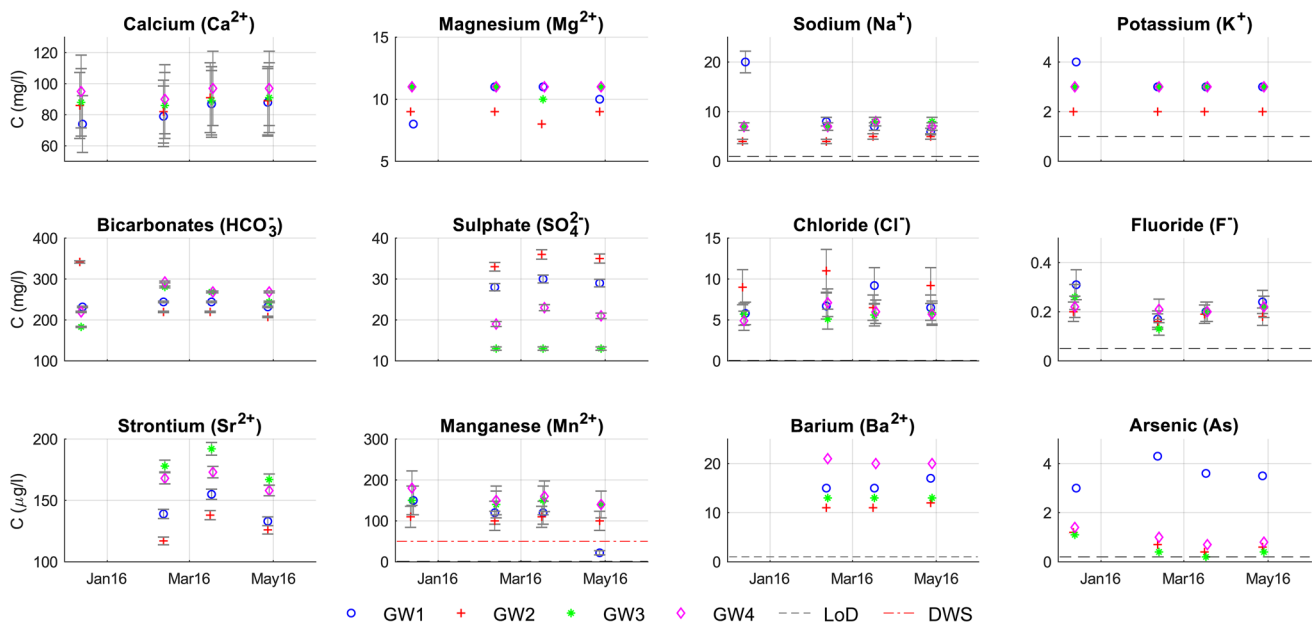
confirmed for bicarbonates since the uncertainties for alkalinity are likely to be underestimated for the reasons discussed in “Uncertainties” section.

Most minor and trace elements have concentrations near or below the detection limit (Table 2 in ESM). Notable exceptions are manganese and strontium (both about 100 times the detection limit), fluoride (concentrations quite variable), barium (concentrations at least 10 times the detection limit) and arsenic (detected in all samples). Traces of boron, chromium, copper, iron, nickel, antimony and selenium were found on some occasions. The remaining elements were systematically below their respective detection limits. Temporal variations of minor and trace elements in significant concentrations are shown in Fig. 4. Considering uncertainties, strontium concentrations show some significant temporal variability. It is interesting to note that the temporal variations for strontium follow the same trend in the four wells and are similar to sulphate concentration variations. Spatial variation is also quite obvious for strontium, barium and arsenic (less obvious for manganese due to larger uncertainty), with at least one well having a different concentration. This confirms the results from the Kruskal–Wallis test (Table 3).

With respect to compliance with Polish drinking water standards, manganese concentrations exceed the standard by two to three times (Fig. 4; Table 2 in ESM). The detection limit for bromide (0.05 mg/l) is five times the Polish drinking water standard (0.01 mg/l). Bromide was detected on several occasions, suggesting that bromide concentrations for the rest of the samples are close to the detection limit and likely above the Polish drinking water standard. All other elements have concentrations below Polish drinking water standards.



**Fig. 3** Stiff diagrams (Stiff 1951) for samples taken during the baseline monitoring. Each line represents one sample



**Fig. 4** Temporal variation in concentration for major elements, and minor and trace elements in significant concentrations. Errors bars: uncertainties calculated by the ‘duplicate’ method with a 95% confidence level. For all duplicate sets with the same concentrations (Mg, K, Ba and As), no uncertainties could be calculated. Limits of

detection (LoD) and drinking water standards (DWS) shown when included within the display of the y-axis. For chloride and manganese, LoD not visible due to their low value (0.01 mg/l and 1 µg/l, respectively)

### Temperature and specific conductivity

The stability of the groundwater chemistry is confirmed by the continuous monitoring of the specific conductivity and temperature. The specific conductivity was fairly constant during the baseline monitoring period and was similar in all the monitoring boreholes (varying between 440 and 500 µS/cm; Fig. 2 in ESM). Similarly, the temperature was also fairly constant over time in all boreholes, ranging from 7.85 to 8.20 °C (Fig. 3 in ESM). This is confirmed by the results of the Kruskal–Wallis test on sub-datasets at three-week intervals (Table 3).

The impact of purging and sampling was clearly visible on the specific conductivity for all monitoring boreholes except GW3 (Fig. 2 in ESM). The effect was variable from borehole to borehole: a systematic decrease by 60–70 µS/cm was observed in GW1, whereas a systematic increase could be seen in GW4. The purging and sampling also affected the temperature records, with a systematic increase by 0.03 to 0.04 °C in GW2 for instance (Fig. 3 in ESM). Although the aquifer appears globally homogeneous in terms of groundwater quality across all monitoring boreholes, these small changes on sampling appear to reflect ‘new’ groundwater being drawn into the well during sampling and illustrate the results from the Kruskal–Wallis test (Table 3). Prior to the pump being switched on, the sensor ‘sees’ a volume of water which may have been resident in the well for some time, or

which has migrated slowly through the well screen under natural head gradients. As soon as the pump is switched on, this volume is rapidly replaced by a new flow of groundwater from the aquifer, migrating under a high induced head gradient (and possibly encompassing a more three-dimensional flow regime—i.e. water from above and below the screened horizon as well as groundwater within the screened horizon).

### Dissolved gases

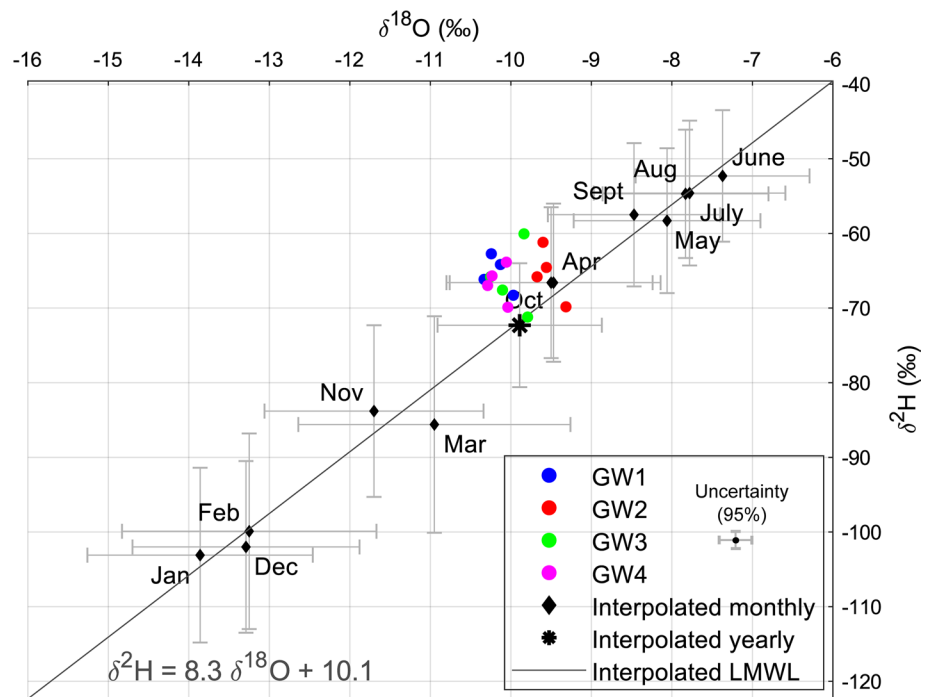
Dissolved gases were not detected during the baseline monitoring, possibly as a result of high detection limits (e.g. 40 µg/l for methane; Table 2 in ESM).

### $\delta^2\text{H}$ and $\delta^{18}\text{O}$ isotope ratios

No GNIP (Global Network of Isotopes in Precipitation; IAEA/WMO 2016) station is present in northern Poland that could provide a comparison for the groundwater data. Instead, interpolated monthly average data (IAEA 2016; Terzer et al. 2013) for the study area were used as a proxy for the local meteoric water line (LMWL; Gunning et al. 2017). The LMWL has a similar equation to the global meteoric water line (GMWL) defined by Craig (1961) (Fig. 5).

Data were in the same range for all monitoring boreholes, with  $\delta^2\text{H}$  ratios varying from  $-71$  to  $-60\text{‰}$  and  $\delta^{18}\text{O}$  ratios from  $-10.3$  to  $-9.3\text{‰}$  (Table 4 in ESM). Significant

**Fig. 5**  $\delta^2\text{H}$  and  $\delta^{18}\text{O}$  isotope ratios measured in groundwater during baseline monitoring. Each coloured point represents one sample (average of duplicates or triplicates). For the sake of clarity, error bars for uncertainty on the measurement are only displayed in the legend. Interpolated monthly and yearly data come from IAEA (2016) and Terzer et al. (2013)



temporal variability was detected by the Kruskal–Wallis test for  $\delta^2\text{H}$  ratios ( $p$  value = 0.01; Table 3) and was confirmed when considering the uncertainties (in particular for GW2 and GW3; Fig. 4 in ESM).  $\delta^{18}\text{O}$  values were slightly less negative in GW2 than in the other boreholes (Table 4 in ESM). This difference appears to be significant (Fig. 4 in ESM) and was detected by the Kruskal–Wallis test ( $p$  value = 0.02; Table 3). With regard to the uncertainties of the yearly interpolated data, all these differences appear to be relatively insignificant (Fig. 5). They are likely to simply result from some natural variation of the recharge and complex recharge pathways.

The data plot close to the LMWL, within the confidence interval of the interpolated LMWL (Fig. 5). The data confirm that the recharge occurs under current climatic conditions, as is expected for aquifers hosted in Quaternary sediments deposited during the last glaciation. The data actually form a small cluster near the interpolated annual mean isotope signature for precipitation, as often observed under temperate climatic conditions (Clark and Fritz 1997).

### Confidence intervals around the mean

Results from the previous statistical tests show some small departure from normality, mostly similar variances and spatial variations between the wells, and limited temporal variability. As a consequence, data from the different wells cannot be pooled together and confidence intervals have to be defined on a per-well basis (intra-well; Helsel and Hirsch 2002; US EPA 2009). Due to the limited number of

data, parametric confidence levels were defined instead of nonparametric intervals, with a confidence level of 95%. In this study, concentrations of naturally occurring inorganic parameters were measured under background conditions. The concentrations and isotope ratios appear to fluctuate around an average value. Therefore, assuming normality would be reasonable for most datasets (US EPA 2009). Nonparametric confidence levels with an appropriate level of confidence can only be defined for temperature and conductivity measurements (Table 4).

The width of the confidence intervals around the mean reflects the presence of temporal variability or the occurrence of one anomalous value that might have been accounted for as an outlier if a larger dataset were available (e.g. Na and Mg in GW1). In that case, this corroborates the departure from normality identified for sodium in GW1 and from equality of variance (Fig. 2 and Table 3). Confidence intervals show their limitation for parameters demonstrating variability and that are close to their detection limit, e.g. arsenic in GW3 with a negative value for its lower limit.

Taking the minimum and maximum values as lower and upper limits, nonparametric confidence intervals have been defined for temperature and conductivity with a confidence level of ~99%. As a result, they are slightly wider than their parametric counterparts. Parametric confidence levels at 99% confidence level (not shown here) are similar to the nonparametric for temperature (in all wells) and conductivity in GW3 and GW4. They are slightly narrower for the conductivity in GW1 and GW2. The lower and upper limits of the nonparametric intervals are the minimum and

**Table 4** Parametric (P) confidence intervals (CI) around the mean with 95% confidence level for different groundwater quality parameters

Parameter	Unit	Uncertainty	CI type	GW1	GW2	GW3	GW4
Calcium	mg/l	$\pm 18\text{--}24$	P (95%)	$83 \pm 13$	$87 \pm 6$	$89 \pm 3$	$95 \pm 5$
Magnesium	mg/l	–	P (95%)	$10 \pm 2.3$	$8.8 \pm 0.8$	$11 \pm 0.8$	11
Sodium	mg/l	$\pm 0.4\text{--}2.2$	P (95%)	$10 \pm 10$	$4.5 \pm 0.9$	$7.5 \pm 0.9$	$7.3 \pm 0.8$
Potassium	mg/l	–	P (95%)	$3.3 \pm 0.8$	2	3	3
Alkalinity	mg/l CaCO <sub>3</sub>	$\pm 1\text{--}2$ (15–28)	P (95%)	$195 \pm 9$	$203 \pm 83$	$200 \pm 57$	$215 \pm 40$
Sulphate	mg/l	$\pm 0.4\text{--}1.2$	P (95%)	$29 \pm 2$	$35 \pm 4$	13	$21 \pm 5$
Chloride	mg/l	$\pm 1.3\text{--}2.6$	P (95%)	$7.1 \pm 2.4$	$8.9 \pm 3.0$	$5.6 \pm 0.6$	$5.9 \pm 1.5$
Strontium	μg/l	$\pm 3\text{--}5$	P (95%)	$142 \pm 29$	$127 \pm 26$	$179 \pm 31$	$166 \pm 19$
Manganese	μg/l	$\pm 5\text{--}42$	P (95%)	$103 \pm 89$	$105 \pm 9$	$145 \pm 9$	$158 \pm 18$
Barium	μg/l	–	P (95%)	$16 \pm 2.9$	$11 \pm 1.5$	13	$20 \pm 1.5$
Fluoride	mg/l	$\pm 0.02\text{--}0.05$	P (95%)	$0.23 \pm 0.10$	$0.18 \pm 0.03$	$0.20 \pm 0.09$	$0.21 \pm 0.02$
Arsenic	μg/l	–	P (95%)	$3.6 \pm 0.9$	$0.7 \pm 0.5$	$0.5 \pm 0.7$	$1.0 \pm 0.5$
δ <sup>2</sup> H	‰	$\pm 3$	P (95%)	$-65 \pm 4$	$-65 \pm 6$	$-66 \pm 8$	$-67 \pm 4$
δ <sup>18</sup> O	‰	$\pm 0.3$	P (95%)	$-10.2 \pm 0.3$	$-9.5 \pm 0.3$	$-10.0 \pm 0.4$	$-10.2 \pm 0.2$
Temperature	°C	$\pm 0.14$	P (95%)	$7.876 \pm 0.003$	$8.114 \pm 0.007$	$8.081 \pm 0.007$	$8.085 \pm 0.004$
			NP (~99%)	7.870–7.880	8.107–8.120	8.070–8.097	8.077–8.090
			P (95%)	$0.474 \pm 0.015$	$0.458 \pm 0.006$	$0.448 \pm 0.003$	$0.468 \pm 0.006$
Conductivity	μS/cm	$\pm 0.006\text{--}0.007$	NP (~99%)	0.432–0.484	0.444–0.468	0.444–0.452	0.456–0.476
			P (95%)	$0.474 \pm 0.015$	$0.458 \pm 0.006$	$0.448 \pm 0.003$	$0.468 \pm 0.006$

Nonparametric (NP) intervals around the mean are given for temperature and conductivity measurements. Unique values indicate that the concentrations were constant during baseline monitoring. Uncertainties at 95% (converted to mg/l) are given for comparison. For alkalinity, theoretical uncertainty is in parentheses (see “[Uncertainties](#)” section). Alkalinity unit is mg/l CaCO<sub>3</sub>

maximum values in the datasets at this level of confidence. With such a limited amount of data, nonparametric intervals are sensitive to the dataset values and are not as robust as the parametric ones.

## Discussion

### Comparison between uncertainty and confidence intervals

The width of the confidence intervals around the mean is compared with the uncertainty on the measurements where both are defined (Table 4). Three cases can be distinguished with the width of confidence intervals similar to, higher than or lower than uncertainties (illustrated in Fig. 6).

The width of the confidence interval is similar to the estimated uncertainties for sodium (except in GW1), chloride (except GW3), fluoride (GW2 and GW4), δ<sup>18</sup>O and specific conductivity (GW2 and GW4). In these cases, the temporal variations reflect the uncertainties on the measurements rather than true temporal variations. Therefore, these parameters do not show any significant temporal variations.

Confidence intervals wider than uncertainties can be explained by the presence of anomalous values, e.g. sodium and manganese in GW1 (Fig. 6). In the case of alkalinity, although the uncertainties might have been underestimated,

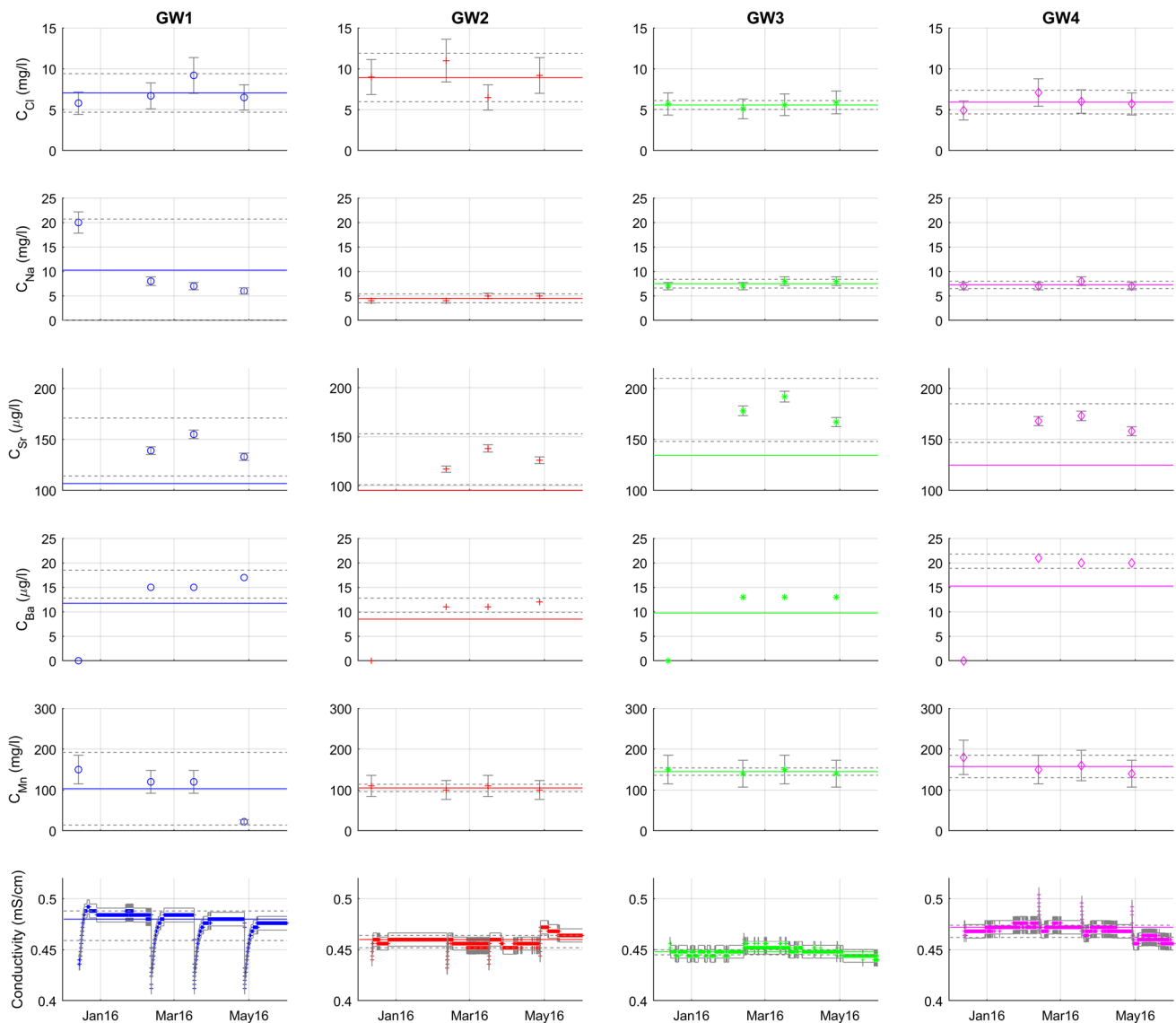
the maximal theoretical uncertainties are still much less than the width of the confidence interval (with the exception of GW1). In other cases (sulphate, fluoride—GW1 and GW4; strontium, δ<sup>2</sup>H and specific conductivity—GW1), it suggests the presence of temporal variations within the wells. This comparison confirms what could be qualitatively inferred from Fig. 4 and Fig. 4 of the ESM.

Uncertainty can be wider than the confidence intervals, e.g. calcium, chloride (GW3), manganese (except GW1), temperature and conductivity (GW3). For calcium and manganese, the uncertainty could be affected by the limited number of duplicates and therefore be overestimated. In particular, for calcium, one pair of duplicates has a large concentration difference. Manganese concentrations are expressed with two significant digits (rounded up to the nearest ten for concentrations above 100 mg/l and to the nearest digit for concentrations below 100 mg/l). The impact on the uncertainty is similar to that described for alkalinity (“[Uncertainties](#)” section). These parameters are essentially constant during baseline monitoring.

### Comparison with other regional studies

Pruszkowska and Malina (2008) defined groundwater background quality levels for the Kashubian Lake District, based on the analyses of groundwater from ~1400 wells between 1990 and 2005. With regard to their study, the Wysin site is





**Fig. 6** Uncertainties on the measurements (grey error bars) and parametric confidence intervals at 95% (grey dashed lines) around the mean (straight line) for a selection of parameters (Cl, Na, Sr, Ba, Mn and specific conductivity) considered as indicators of potential con-

tamination by shale gas activities. For conductivity, the median (all data) is showed instead of the mean. Same colour scheme as in previous figures

located in the southern part of the area under consideration. Background levels of the ‘deeper aquifers’ group are used for comparison (Table 5). The concentration range observed at the Wysin site during the baseline monitoring study is mostly within background levels defined by Pruszkowska and Malina (2008). Calcium concentrations at Wysin tend towards the upper limit of the background concentrations, with concentrations in GW4 sitting just outside the limit (Fig. 7). The range of calcium concentrations in Pruszkowska and Malina (2008)’s study extends to 185 mg/l. Only the range of sodium concentrations is available, with concentrations up to 16 mg/l. The 20 mg/l value in GW1 is certainly unusual and appears to be an outlier.

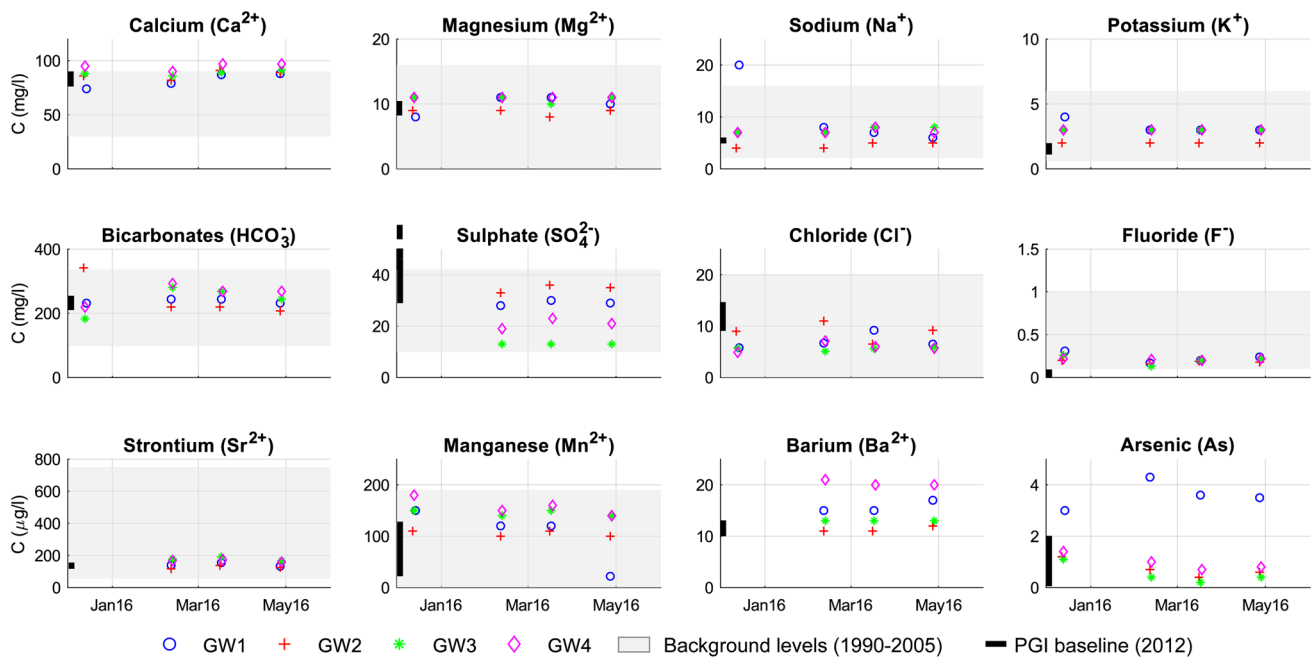
Isotope data from the Kashubian Lake District are available in Pruszkowska and Malina (2008), for seven wells located 30 km north of the Wysin site. Three additional analysis results are presented in Kachnic and Kachnic (2010), in wells located up to 50 km south of the Wysin site. The  $\delta^2\text{H}$  and  $\delta^{18}\text{O}$  isotope signatures observed during the baseline monitoring are within the range measured by Pruszkowska and Malina (2008). The range from Kachnic and Kachnic (2010) is narrower due to the limited number of data points, although they are from a wider area, and is within the range observed in this study (Table 5).

The  $\delta^2\text{H}$  and  $\delta^{18}\text{O}$  signatures at the Wysin site are not depleted with respect to the weighted mean for

**Table 5** Comparison of baseline monitoring results from this study with background levels published for the region (Pruszkowska and Malina 2008). Additional isotope data from Kachnic and Kachnic (2010)

Parameter	Unit	Range (this study)	Background levels (Pruszkowska and Malina 2008)	Range (Kachnic and Kachnic 2010)
Calcium	mg/l	74–97	30–90	—
Magnesium	mg/l	8–11	0–16	—
Sodium	mg/l	4–20	2.1–16.0*	—
Potassium	mg/l	2–4	0.6–6.0	—
Alkalinity	mg/l CaCO <sub>3</sub>	150–280	80–275	—
Sulphate	mg/l	13–36	10–42	—
Chloride	mg/l	4.9–11	0–20	—
Strontium	µg/l	117–192	56–750*	—
Manganese	µg/l	22–180	0–190	—
Fluoride	mg/l	0.13–0.31	0.1–1*	—
δ <sup>2</sup> H	‰	−72 to −60	−71 to −57*	−71.6 to −69.9*
δ <sup>18</sup> O	‰	−10.3 to −9.4	−10.7 to −7.6*	−9.98 to −9.83*

\*No background provided: range provided instead

**Fig. 7** Comparison of the baseline data from this study with natural regional background concentrations (grey shaded areas) from Pruszkowska and Malina (2008) and 2012 baseline data (black line) from Koniecznyńska et al. (2014)

precipitation, indicating that snowmelt does not form an essential part of the recharge, contrary to northern parts of the BAB (Raidla et al. 2016). Such depletion would also be enhanced by low recharge rates due to evapotranspiration processes in the summer (recharge pattern typical of temperate climates). The absence of depletion with respect to the weighted mean for precipitation suggests complex recharge processes due to the semi-confinement of the aquifer, the depth of the sampling points, the distance from the potential recharge area and the groundwater

velocity resulting in a good mixing of groundwater (Clark and Fritz 1997).

### Baseline conditions at Wysin

The PGI-NRI carried out a baseline monitoring programme in September 2012, prior to any drilling at the shale gas site (Koniecznyńska et al. 2014). As part of this programme, 12 water samples were collected (location of the sampling points shown in Fig. 1). The dug wells and the manual

soundings tap into the shallow unconfined/perched aquifer. Of the two springs, one is considered to be representative of the groundwater quality of the main aquifer (Konieczńska et al. 2014). For comparison with this study, only the samples for the drilled boreholes and the above-mentioned spring are considered to be related to the main aquifer level. Historical data are available for one of the drilled boreholes (from 1976 at the time of the drilling). Analysis data are also available for a new well drilled nearby in 2014. Data are compiled in Table 6.

Major ion concentrations in this study are in the same range as those from the PGI baseline (Fig. 7) and the historical data. Minor and trace elements with concentrations largely above their respective limits of detection are also in the same range, i.e. strontium, manganese and barium (Fig. 7). For those close to the limits of detection, concentrations are highly variable between the two baselines (Table 6). It is not surprising since low concentrations are quite variable within the same well. Additionally, the use

of different laboratories with different limits of detection is likely to impact more on low concentrations near the limit of detection.

## Conclusions and recommendations

This paper presents the results from one of the first baseline monitoring programmes in the context of shale gas development in Europe. The monitoring network consists of four dedicated boreholes. Due to project time constraints, the baseline monitoring only lasted 6 months during which four sampling events took place.

Statistics used to analyse the data revealed their limitations for application to such small datasets, in particular for the use of robust nonparametric tests. Some limitations can be overcome by quantifying the uncertainty stemming from analysis and sampling. Comparison of the uncertainties with confidence intervals around the mean helps to better identify

**Table 6** Comparison of baseline monitoring data from this study with those from the baseline monitoring carried out in 2012 by the Polish Geological Institute (Konieczńska et al. 2014). Additional analysis results available for boreholes in Stry Wiek are included (courtesy of PGI-NRI)

Parameter	Unit	Range (this study)	Range (Konieczńska et al. 2014)	Historical data	
				1976	2014
Calcium	mg/l	74–97	78–91	< 90*	< 148*
Magnesium	mg/l	8–11	8.3–10.2	–	–
Sodium	mg/l	4–20	4.9–5.9	–	11.5
Potassium	mg/l	2–4	1.2–2.0	–	3.72
Alkalinity	mg/l CaCO <sub>3</sub>	150–280	172–209	200	265
Sulphate	mg/l	13–36	27.6–70.7	30.4	132
Chloride	mg/l	4.9–11	8.93–14.9	16	16.5
Strontium	µg/l	117–192	112–149	–	–
Manganese	µg/l	22–180	20–128	< LoD	84
Barium	µg/l	11–21	10–13	–	–
Fluoride	mg/l	0.13–0.31	< 0.10	–	–
Arsenic	µg/l	0.2–4.3	< 2	–	–
Iron	mg/l	< 0.01–0.1	< 0.01–1.15	0.2	0.49
Aluminium	mg/l	< 0.01–0.05	0.008–0.013	–	–
Boron	mg/l	< 0.1–0.8	0.01–0.02	–	–
Bromide	mg/l	< 0.05–0.09	< 0.10	–	–
Cadmium	µg/l	< 0.02	< 0.05	–	–
Chromium	µg/l	< 1.0–8	< 3	–	–
Copper	µg/l	< 0.5–0.6	0.33–0.96	–	–
Mercury	µg/l	< 0.05	< 0.3	–	–
Lithium	mg/l	< 0.1	0.0031–0.0066	–	–
Nickel	µg/l	< 1.0–6	< 0.5	–	–
Lead	µg/l	< 0.3	< 0.05–0.07	–	–
Antimony	µg/l	< 1.0–2	< 0.05	–	–
Selenium	µg/l	< 0.5–1.3	< 2	–	–
Methane	mg/l	< 0.04	< 0.0001–0.0016	–	–
Conductivity	µS/cm	440–490	433–534	–	711

\*Only hardness is available and is converted considering only calcium to give an upper limit  
LoD limit of detection

spatial and temporal variations for small datasets. From a statistical point of view, it is recommended to start baseline monitoring as early as possible in order to allow the collection of at least six statistically independent samples evenly distributed during the course of a year. This will enable the use of statistics as an unbiased and robust method to better characterise baseline conditions (e.g. temporal variability) following the framework outlined in this paper. It is also recommended that a robust quality control scheme is designed and implemented at the outset of the baseline monitoring programme in order to quantify data uncertainty. Like small datasets, large datasets from downhole probes present some challenges for the use of statistics. This is an area that requires further investigation considering their widespread use and the possibility of using remote sensing capabilities for live monitoring. Further work is also required on how to identify quality changes from baseline conditions (for both types of datasets), not only from a scientific point of view but from the regulators' perspective.

The groundwater composition around the shale gas pad is typical of recently infiltrated water hosted in Quaternary sediments (Ca–HCO<sub>3</sub> water type). The groundwater is characterised by a low mineral content and stable temperature. The groundwater quality is good, complying with most requirements in terms of drinking water quality, except for manganese and bromide. The groundwater composition is relatively homogeneous around the drilling site and is similar to background levels in other aquifers in the region. The baseline groundwater quality at the monitoring boreholes is therefore representative of the regional groundwater quality, and the monitoring boreholes are deemed suitable for their task. However, median concentrations are statistically different between the boreholes for most parameters. Therefore, with the aim of detecting future potential contamination by shale gas development or other anthropogenic activities, baseline conditions should be defined for each borehole for these parameters. The groundwater composition showed limited temporal variability for most parameters which is accounted for when defining the confidence intervals around the mean. For parameters with no detected spatial and temporal variations, the datasets from all boreholes would be pooled together to obtain a larger dataset and define more robust baseline conditions at site level.

Comparison with the baseline carried out by the PGI-NRI in 2012 shows similar groundwater concentrations. Since different wells from this study were sampled only once, it is difficult to reach conclusions on any impacts from the shale gas activities prior to hydraulic fracturing itself (i.e. from the drilling of vertical and horizontal wells in the period 2012–2015). While the focus of this study has been on inorganic constituents (apart from dissolved gases), the PGI-NRI baseline analysed the samples for a range of organic contaminants. Some organic contaminants relevant to the shale

gas industry should be included during baseline monitoring. The analytical suite could be reduced during operational monitoring to reduce costs and be expanded again if any suspicion arises.

**Acknowledgements** This work has been carried out as part of the SHEER project, which has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 640896. The authors would like to acknowledge the help and support given to the project by the late Professor Paul Younger. The authors would also like to thank Prof. Adrian Boyce and Alison McDonald for running isotope analyses at the ICSF (NERC Facility contract F14/G6/11/01) within the SUERC laboratories.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Bair ES, Freeman DC, Senko JM (2010) Subsurface gas invasion Bainbridge Township, Geauga County, Ohio. Ohio Department of Natural Resources, Columbus, OH. <https://oilandgas.ohiodnr.gov/portals/oilgas/pdf/bainbridge/DMRM%200%20Title%20Page,%20Preface,%20Acknowledgements.pdf>. Accessed Oct. 2017
- Bartels R (1982) The rank version of von Neumann's ratio test for randomness. *J Am Stat Assoc* 77:40–46. <https://doi.org/10.2307/2287767>
- Bell RA, Darling WG, Ward RS, Basava-Reddi L, Halwa L, Manamsa K, Dochartaigh BEÓ (2017) A baseline survey of dissolved methane in aquifers of Great Britain. *Sci Total Environ* 601:1803–1813. <https://doi.org/10.1016/j.scitotenv.2017.05.191>
- Brownfield ME, Schenk CJ, Charpentier RR, Klett TR, Pitman JK, Tennyson ME, Gaswirth SB, Mercier TJ, Le PA, Leathers HM (2015) Assessment of undiscovered conventional and continuous oil and gas resources of the Baltic Depression Province. US Geological Survey. <https://doi.org/10.3133/fs20153015>. <https://pubs.er.usgs.gov/publication/fs20153015>. Accessed June 2018
- Cantoni R (2018) Second Galicia? Poland's shale gas rush through historical lenses. *Geol Soc Lond Spec Publ* 465:SP465–416
- Chou CJ (2004) Groundwater monitoring: statistical methods for testing special background conditions. In: Wiersma GB (ed) *Environmental monitoring*. CRC Press, Boca Roca, pp 239–255
- Clark ID, Fritz P (1997) *Environmental isotopes in hydrogeology*. CRC Press LLC, Boca Raton
- Craig H (1961) Isotopic variations in meteoric waters. *Science* 133(3465):1702–1703. <https://doi.org/10.1126/science.133.3465.1702>
- Darrah TH, Vengosh A, Jackson RB, Warner NR, Poreda RJ (2014) Noble gases identify the mechanisms of fugitive gas contamination in drinking-water wells overlying the Marcellus and

- Barnett Shales. *Proc Natl Acad Sci* 111:14076–14081. <https://doi.org/10.1073/pnas.1322107111>
- Donnelly T, Waldron S, Tait A, Dougans J, Bearhop S (2001) Hydrogen isotope analysis of natural abundance and deuterium-enriched waters by reduction over chromium on-line to a dynamic dual inlet isotope-ratio mass spectrometer. *Rapid Commun Mass Spectrom* 15:1297–1303
- EC (1998) Council directive 98/83/EC of 3 November 1998 on the quality of water intended for human consumption. *Off J Eur Communities* 41:32–54
- EC (2018) Proposal for a directive of the European Parliament and of the Council on the quality of water intended for human consumption (recast). European Commission, Brussels, Belgium. <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52017PC0753&from=EN>. Accessed May 2018
- EEA (2014) Interactive maps. European Environmental Agency. <http://maps.eea.europa.eu/EEABasicviewer/v3/?appid=d649ddee9e145559a650ac6ed714df&embed=true>. Accessed Nov. 2018
- Global Network of Isotopes in Precipitation. The GNIP database (2016) <http://www-naweb.iaea.org/naweb/ih/index.html>
- Gordalla BC, Ewers U, Frimmel FH (2013) Hydraulic fracturing: a toxicological threat for groundwater and drinking-water? *Environ Earth Sci* 70:3875–3893. <https://doi.org/10.1007/s12665-013-2672-9>
- Grøn C, Hansen JB, Magnusson B, Nordbotten A, Krysell M, Andersen KJ, Lund U (2007) Uncertainty from sampling—a NORDTEST handbook for sampling planners on sampling quality assurance and uncertainty estimation. Nordic Innovation Centre, Norway. <http://demarcheisoi17025.com/documents/Uncertainty%20from%20Sampling%20-%20NORDTEST%20Report%20-%20TR604.pdf>. Accessed Dec 2017
- GT (2015) Dokumentacja geologiczna z wykonania 7 otworów badawczych w rejonie miejscowości Wysin i Stary Wiec [Geological documentation from drilling seven research boreholes in the region of Wysin and Stary Wiec]. Geofizyka Toruń, Toruń
- Gunning A, Isherwood C, Montcoudiol N, Kelly T (2017) D5.1 Hydrogeological models to assess the potential for groundwater impairment for a selected ‘shale analogue’ site. RSKW Ltd & University of Glasgow. <http://www.sheerproject.eu/images/deliverables/SHEER-Deliverable-5.1.pdf>. Accessed April 2018
- Harkness JS, Darrah TH, Warner NR, Whyte CJ, Moore MT, Millot R, Kloppmann W, Jackson RB, Vengosh A (2017) The geochemistry of naturally occurring methane and saline groundwater in an area of unconventional shale gas development. *Geochim Cosmochim Acta* 208:302–334. <https://doi.org/10.1016/j.gca.2017.03.039>
- Helsel D, Hirsch R (2002) Statistical methods in water resources. US Geological Survey. <https://pubs.usgs.gov/twri/twri4a3/pdf/twri4a3-new.pdf>. Accessed June 2018
- Howarth RW, Ingraffea A, Engelder T (2011) Should fracking stop? *Nature* 477:271. <https://doi.org/10.1038/477271a>
- Humez P, Mayer B, Nightingale M, Ing J, Becker V, Jones D, Lam V (2016) An 8-year record of gas geochemistry and isotopic composition of methane during baseline sampling at a groundwater observation well in Alberta (Canada). *Hydrogeol J* 24:109–122
- Huw C, Leo E, Peter S, Peter T (2014) Felt seismicity associated with shale gas hydraulic fracturing: the first documented example in Europe. *Geophys Res Lett* 41:8308–8314. <https://doi.org/10.1002/2014GL062047>
- IAEA (2016) RCWIP (Regionalized cluster-based water isotope prediction) Model—gridded precipitation  $\delta^{18}\text{O}$ ,  $\delta^2\text{H}$  and  $\delta^3\text{H}$  isoscape data. International Atomic Energy Agency. <http://www-naweb.iaea.org/naweb/ih/index.html>. Accessed Sept 2017
- Jackson RE, Heagle DJ (2016) Sampling domestic/farm wells for baseline groundwater quality and fugitive gas. *Hydrogeol J* 24:269–272. <https://doi.org/10.1007/s10040-016-1369-z>
- Jackson RB, Vengosh A, Darrah TH, Warner NR, Down A, Poreda RJ, Osborn SG, Zhao K, Karr JD (2013a) Increased stray gas abundance in a subset of drinking water wells near Marcellus shale gas extraction. *Proc Natl Acad Sci* 110:11250–11255. <https://doi.org/10.1073/pnas.1221635110>
- Jackson RE, Gorody AW, Mayer B, Roy JW, Ryan MC, Van Stempvoort DR (2013b) Groundwater protection and unconventional gas extraction: the critical need for field-based hydrogeological research. *Groundwater* 51:488–510
- Jaworska-Szulc B (2009) Groundwater flow modelling of multi-aquifer systems for regional resources evaluation: the Gdansk hydrogeological system, Poland. *Hydrogeol J* 17:1521–1542
- JCGM (2008) Evaluation of measurement data—guide to the expression of uncertainty in measurement. Joint Committee for Guides in Metrology (BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML). [https://www.bipm.org/utls/common/documents/jcgml/JCGM\\_100\\_2008\\_E.pdf](https://www.bipm.org/utls/common/documents/jcgml/JCGM_100_2008_E.pdf). Accessed June 2018
- Kachnic J, Kachnic M (2010) Hydrodynamic and hydrochemical conditions in the Wda and Maława River catchments (NW Poland). *Geol Q* 50:447–456
- Kahrilas GA, Blotvogel J, Stewart PS, Borch T (2015) Biocides in hydraulic fracturing fluids: a critical review of their usage, mobility, degradation, and toxicity. *Environ Sci Technol* 49:16–32. <https://doi.org/10.1021/es503724k>
- Koniecznyńska M, Woźnicka M, Antolak O, Janica R, Lichtarski G, Nidental M, Otwinowski J, Starzycka A, Stec B, Wróbel G, Potrykus R, Gdaniec Rohde B, Włodarski T (2011) Environmental aspects of hydraulic fracturing treatment performed on the Łebień LE-2H well. Final report. Polish Geological Institute, National Research Institute, Warsaw, Poland. <https://www.pgi.gov.pl/en/dokumenty-pig-pib-all/kopalnia-wiedzy/gaz-lupkowy/786-the-lebien-report/file.html>
- Koniecznyńska M, Janica R, Fajfer J, Felter A, Frydel J, Głuszyński A, Gryczko-Gostyńska A, Jarosiński M, Kijewska S, Lichtarski G, Lipińska O, Pergół S, Podhalańska T, Roman M, Rosowiecka O, Stec B, Woźnicka M, Dzieńiewicz M, Konopka E, Kotarba M, Lewkiewicz-Kołyś A, Macuda J, Nagy S, Sechman H, Bernaciak M, Grzelak W, Janicki W, Korkosz A, Kozak K, Kudlak B, Męcik M, Zabiegała B (2014) Określenie zakresu oddziaływania procesu poszukiwania i eksploatacji niekonwencjonalnych złóż węglowodorów na środowisko, z uwzględnieniem terminów poszczególnych prac, infrastruktury podziemnej i przesyłowej, a w szczególności określenia charakteru i zakresu oddziaływania [The environment and shale gas exploration—results of studies on the soil-water environment, ambient air, acoustic climate, process fluids and wastes]. Państwowy Instytut Geologiczny–Państwowy Instytut Badawczy, Akademia Górniczo-Hutnicza w Krakowie & Politechnika Gdańska [Polish Geological Institute–National Research Institute, AGH University of Science and Technology in Krakow & Gdańsk University of Technology], Warsaw, Poland
- Koniecznyńska M, Adamczak-Biały T, Brodecki A, Brzezińska A, Janica R, Dziekan-Kamińska E, Fajfer J, Feldman-Olszewska A, Felter A, Frydel J, Głuszyński A, Gryczko-Gostyńska A, Jarosiński M, Joźwiak K, Kordalski Z, Kowalewski T, Kijewska S, Lichtarski G, Lidzbarski J, Lipińska O, Mikołajków J, Nidental M, Otwinowski J, Pasierowska B, Pergół S, Podhalańska T, Roman M, Rosowiecka O, Sobień K, Starzycka A, Stec B, Śliwiński Ł, Waksmundzka M, Woźnicka M, Dzieńiewicz M, Guzy P, Izidor G, Konopka E, Kotarba M, Kowalski T, Lewkiewicz-Kołyś A, Macuda J, Nagy S, Sechman H, Bernaciak M, Grzelak W, Janicki W, Korkosz A, Kozak K, Kudlak B, Męcik M, Zabiegała B (2015) The environment and shale gas exploration—results of studies on the soil-water environment, ambient air, acoustic climate, process fluids and wastes. Directorate General for Environmental Protection,



- Warsaw, Poland. [http://www.gdos.gov.pl/files/artykuly/38173/The\\_environment\\_and\\_shale\\_gas\\_exploration\\_RESULTS\\_OF\\_STUDIES\\_ON\\_THE\\_SOIL\\_WATER\\_ENVIRONMENT\\_AMBIENT\\_AIR\\_ACOUSTIC\\_CLIMATE\\_PROCESS\\_FLUID\\_S\\_AND\\_WASTES.pdf](http://www.gdos.gov.pl/files/artykuly/38173/The_environment_and_shale_gas_exploration_RESULTS_OF_STUDIES_ON_THE_SOIL_WATER_ENVIRONMENT_AMBIENT_AIR_ACOUSTIC_CLIMATE_PROCESS_FLUID_S_AND_WASTES.pdf). Accessed Nov 2017
- Kruskal WH, Wallis WA (1952) Use of ranks in one-criterion variance analysis. *J Am Stat Assoc* 47:583–621. <https://doi.org/10.1080/01621459.1952.10483441>
- Lefebvre R (2017) Mechanisms leading to potential impacts of shale gas development on groundwater quality. *Wiley Interdiscip Rev Water* 4:e1188. <https://doi.org/10.1002/wat2.1188>
- Levene H (1961) Robust tests for equality of variances. In: Olkin I, Ghurye SG, Hoeffding W, Madow WG, Mann HB (eds) Contributions to probability and statistics: essays in honor of Harold Hotelling. Stanford University Press, Palo Alto, pp 279–292
- López-Comino JA, Cesca S, Jarosławski J, Montcoudiol N, Heimann S, Dahm T, Lasocki S, Gunning A, Capuano P, Ellsworth WL (2018) Induced seismicity response of hydraulic fracturing: results of a multidisciplinary monitoring at the Wysin site, Poland. *Sci Rep* 8:8653. <https://doi.org/10.1038/s41598-018-26970-9>
- Luek JL, Gonsior M (2017) Organic compounds in hydraulic fracturing fluids and wastewaters: a review. *Water Res* 123:536–548. <https://doi.org/10.1016/j.watres.2017.07.012>
- Mair R, Bickle M, Goodman D, Koppelman B, Roberts J, Selley R, Shipton Z, Thomas H, Walker A, Woods E, Younger PL (2012) Shale gas extraction in the UK: a review of hydraulic fracturing. Royal Society and Royal Academy of Engineering, London, UK. <https://www.raeng.org.uk/publications/reports/shale-gas-extraction-in-the-uk>. Accessed June 2018
- Makos M (2014) Activities of PGNiG in exploration of unconventional hydrocarbon resources in Poland. Institute of Geological Sciences of Polish Academy of Sciences (ING PAN), Warsaw, Poland. [http://www.ing.pan.pl/Atlab/Presentations/Sci\\_4\\_Ind\\_2014/Presentations/06\\_presentation\\_PGNiG.pdf](http://www.ing.pan.pl/Atlab/Presentations/Sci_4_Ind_2014/Presentations/06_presentation_PGNiG.pdf). Accessed May 2018
- Montcoudiol N, Banks D, Isherwood C, Gunning A (2018) D5.2 Impact of well construction and fracture stimulation on baseline hydrogeological conditions and on drinking water aquifers. University of Glasgow & RSKW Ltd. <https://doi.org/10.1016/j.egypr.2017.08.083> Accessed May 2018
- Moritz A, Hélie J-F, Pinti DL, Larocque M, Barnette D, Retailliau S, Lefebvre R, Gélinais Y (2015) Methane baseline concentrations and sources in shallow aquifers from the shale gas-prone region of the St. Lawrence Lowlands (Quebec, Canada). *Environ Sci Technol* 49:4765–4771. <https://doi.org/10.1021/acs.est.5b00443>
- Nelson ST (2000) A simple, practical methodology for routine VSMOW/SLAP normalization of water samples analyzed by continuous flow methods. *Rapid Commun Mass Spectrometry* 14:1044–1046. [https://doi.org/10.1002/1097-0231\(20000630\)14:12%3C1044:AID-RCM987%3E3.0.CO;2-3](https://doi.org/10.1002/1097-0231(20000630)14:12%3C1044:AID-RCM987%3E3.0.CO;2-3)
- OGI (n.d.) Environmental impact of the exploration and recognition of the crude oil and natural gas deposits in the concession area 1/2011/p STARA KISZEWA [in Polish]. Oil and Gas Institute, Krakow, Poland
- PGI-NRI (2012) Assessment of shale gas and shale oil resources of the Lower Paleozoic Baltic-Podlasie-Lublin Basin in Poland. Polish Geological Institute–National Research Institute, Warsaw, Poland. <https://www.pgi.gov.pl/en/dokumenty-pig-pib-all/aktualnosc-i-2012/zasoby-gazu/769-raport-en/file.html>. Accessed May 2018
- Pólya G (1920) Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentenproblem [On the central limit theorem of probability calculation and the problem of moments]. *Math Z* 8:171–181. <https://doi.org/10.1007/BF01206525>
- Pruszkowska M, Malina G (2008) Hydrogeochemistry and vulnerability of groundwater in the moraine upland aquifers of the Gdańsk region (Northern Poland). *Geol Q* 52:291–300
- Pyssa J (2017) The influence of shale gas mining activities on the natural environment in Poland. In: Energy, environment and material systems (EEMS), Polanica Zdrój, Poland, 13–15 Sept. 2017. E3S Web of Conferences. EDP Sciences. <https://doi.org/10.1051/e3sconf/20171902024>
- Raidla V, Kern Z, Pärn J, Babre A, Erg K, Ivask J, Kalvāns A, Kohān B, Lelgus M, Martma T, Mokrik R, Popovs K, Vaikmäe R (2016) A  $\delta^{18}\text{O}$  isoscape for the shallow groundwater in the Baltic Artesian Basin. *J Hydrol* 542:254–267. <https://doi.org/10.1016/j.jhydrol.2016.09.004>
- Ramsey MH, Ellison SLR (eds) (2007) Measurement uncertainty arising from sampling: a guide to methods and approaches. Eurachem/EUROLAB/CITAC/Nordtest/AMC Guide
- Rhodes AL, Horton NJ (2015) Establishing baseline water quality for household wells within the Marcellus Shale gas region, Susquehanna County, Pennsylvania, U.S.A. *Appl Geochem* 60:14–28. <https://doi.org/10.1016/j.apgeochem.2015.03.004>
- Sadurski A (1986) Hydrogeological evolution of the upper cretaceous artesian basin of the Gdansk region. *Ann Soc Geol Pol* 56:143–161
- Schloemer S, Elbracht J, Blumenberg M, Illing CJ (2016) Distribution and origin of dissolved methane, ethane and propane in shallow groundwater of Lower Saxony, Germany. *Appl Geochem* 67:118–132. <https://doi.org/10.1016/j.apgeochem.2016.02.005>
- Schnoor JL (2012) Shale gas and hydrofracturing. *Environ Sci Technol* 46:4686. <https://doi.org/10.1021/es3011767>
- Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). *Biometrika* 52:591–611. <https://doi.org/10.1093/biomet/52.3-4.591>
- Siegel DI, Smith B, Perry E, Bothun R, Hollingsworth M (2015) Pre-drilling water-quality data of groundwater prior to shale gas drilling in the Appalachian Basin: analysis of the Chesapeake Energy Corporation dataset. *Appl Geochem* 63:37–57. <https://doi.org/10.1016/j.apgeochem.2015.06.013>
- Sloto RA (2014) Baseline groundwater quality from 34 wells in Wayne County, Pennsylvania, 2011 and 2013. Reston, VA. <https://doi.org/10.3133/ofr20141116>. <http://pubs.er.usgs.gov/publication/ofr20141116>
- Smedley PL, Ward RS, Allen G, Baptie BJ, Daraktchieva Z, Jones DG, Jordan CJ, Purvis RM, Cigna F (2015) Site selection strategy for environmental monitoring in connection with shale-gas exploration: Vale of Pickering, Yorkshire and Fylde, Lancashire. British Geological Survey, Keyworth, UK. <http://nora.nerc.ac.uk/id/eprint/512628/>
- Stiff HA Jr (1951) The interpretation of chemical water analysis by means of patterns. *J Pet Technol* 3:15–17. <https://doi.org/10.2118/951376-G>
- Stringfellow WT, Domen JK, Camarillo MK, Sandelin WL, Borglin S (2014) Physical, chemical, and biological characteristics of compounds used in hydraulic fracturing. *J Hazardous Mater* 275:37–54. <https://doi.org/10.1016/j.jhazmat.2014.04.040>
- Terzer S, Wassenaar L, Araguás-Araguás L, Aggarwal P (2013) Global isoscapes for  $\delta^{18}\text{O}$  and  $\delta^2\text{H}$  in precipitation: improved prediction using regionalized climatic regression models. *Hydrol Earth Syst Sci* 17:4713–4728. <https://doi.org/10.5194/hess-17-1-2013>
- Tukey JW (1977) Exploratory data analysis. Addison-Wesley Series in behavioral science: quantitative methods. Addison-Wesley Publishing Company, Boston
- US EPA (2009) Statistical analysis of groundwater monitoring data at RCRA facilities: Unified guidance. U.S. Environmental Protection Agency. <https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P10055GQ.TXT>. Accessed May 2018
- US EIA/ARI (2013) Technically recoverable shale oil and shale gas resources: an assessment of 137 shale formations in 41 countries outside the United States. US Energy Information Agency & Advanced Research Institute, Washington, USA. <https://www.eia>

- [gov/analysis/studies/worldshalegas/pdf/overview.pdf](https://www.epa.gov/analysis/studies/worldshalegas/pdf/overview.pdf). Accessed June 2018
- US EPA (2015) Retrospective case study in Killdeer, North Dakota—study of the potential impacts of hydraulic fracturing on drinking water resources. Office of Research and Development, Washington, DC. [https://www.epa.gov/sites/production/files/2015-06/documents/killdeer\\_report\\_508\\_km\\_1.pdf](https://www.epa.gov/sites/production/files/2015-06/documents/killdeer_report_508_km_1.pdf). Accessed May 2018
- US EPA (2016) Hydraulic fracturing for oil and gas: Impacts from the hydraulic fracturing water cycle on drinking water resources in the United States. Office of Research and Development, Washington, DC. <https://cfpub.epa.gov/ncea/hfstudy/recordisplay.cfm?deid=332990>
- Vidic RD, Brantley SL, Vandenbossche JM, Yoxtheimer D, Abad JD (2013) Impact of shale gas development on regional water quality. *Science*. <https://doi.org/10.1126/science.1235009>
- Virbulis J, Bethers U, Saks T, Sennikovs J, Timuhins A (2013) Hydrogeological model of the Baltic Artesian Basin. *Hydrogeol J* 21:845–862
- Williams L, Macnaghten P, Davies R, Curtis S (2015) Framing ‘fracking’: Exploring public perceptions of hydraulic fracturing in the United Kingdom. *Public understanding of science* (Bristol, England). <https://doi.org/10.1177/0963662515595159>
- Witczak S, Bronders J, Kania J, Kmiecik E, Rozanski K, Szczepanska J (2006) Deliverable 16: summary guidance and recommendations on sampling, measuring and quality assurance. AGH - University of Science and Technology, Krakow, Poland. [https://hydrologie.org/BIB/Publ\\_UNESCO/SOG\\_BRIDGE/Deliverables/WP3/D16.pdf](https://hydrologie.org/BIB/Publ_UNESCO/SOG_BRIDGE/Deliverables/WP3/D16.pdf). Accessed Dec. 2017